

# Evolutionary position of breviate amoebae and the primary eukaryote divergence

Marianne A. Minge<sup>1</sup>, Jeffrey D. Silberman<sup>2</sup>, Russell J. S. Orr<sup>3</sup>,  
Thomas Cavalier-Smith<sup>4</sup>, Kamran Shalchian-Tabrizi<sup>3</sup>, Fabien Burki<sup>5</sup>,  
Åsmund Skjæveland<sup>3</sup> and Kjetill S. Jakobsen<sup>1,\*</sup>

<sup>1</sup>Department of Biology, Centre for Ecological and Evolutionary Synthesis, University of Oslo, 0316 Oslo, Norway

<sup>2</sup>Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA

<sup>3</sup>Department of Biology, Microbial Evolution Research Group, University of Oslo, 0316 Oslo, Norway

<sup>4</sup>Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

<sup>5</sup>Department of Zoology and Animal Biology, Molecular Systematics Group, University of Geneva, 1224 Chêne-Bougeries, Switzerland

Integration of ultrastructural and molecular sequence data has revealed six supergroups of eukaryote organisms (excavates, Rhizaria, chromalveolates, Plantae, Amoebozoa and opisthokonts), and the root of the eukaryote evolutionary tree is suggested to lie between unikonts (Amoebozoa, opisthokonts) and bikonts (the other supergroups). However, some smaller lineages remain of uncertain affinity. One of these unassigned taxa is the anaerobic, free-living, amoeboid flagellate *Breviata anathema*, which is of key significance as it is unclear whether it is a unikont (i.e. possibly the deepest branching amoebozoan) or a bikont. To establish its evolutionary position, we sequenced thousands of *Breviata* genes and calculated trees using 78 protein sequences. Our trees and specific substitutions in the 18S RNA sequence indicate that *Breviata* is related to other Amoebozoa, thereby significantly increasing the cellular diversity of this phylum and establishing *Breviata* as a deep-branching unikont. We discuss the implications of these results for the ancestral state of Amoebozoa and eukaryotes generally, demonstrating that phylogenomics of phylogenetically ‘nomadic’ species can elucidate key questions in eukaryote evolution. Furthermore, mitochondrial genes among the *Breviata* ESTs demonstrate that *Breviata* probably contains a modified anaerobic mitochondrion. With these findings, remnants of mitochondria have been detected in all putatively deep-branching amitochondriate organisms.

**Keywords:** *Breviata anathema*; bikont; unikont; amoebozoa; excavates; phylogenomics

## 1. INTRODUCTION

Almost all the millions of eukaryote species belong to only six recognized supergroups of organisms (Baldauf 2003; Keeling 2004; Simpson & Roger 2004; Keeling *et al.* 2005). Recent molecular and cellular evidence suggests that these in turn may comprise just two superclades: unikonts and bikonts (Stechmann & Cavalier-Smith 2003; Richards & Cavalier-Smith 2005). The exclusively heterotrophic unikont eukaryotes comprise opisthokonts (animals, fungi and immediate unicellular relatives) and Amoebozoa (amoebae with broad pseudopods and slime moulds), while the bikonts comprise photosynthetic Plantae, chromalveolates (chromophyte algae and their non-photosynthetic descendants, e.g. ciliate and sporozoan protozoa) and two diverse groups of mainly heterotrophic protozoa (excavates, predominantly flagellates with rigid cell cortex and a specialized feeding groove, and Rhizaria, mostly soft-surfaced cells with elaborate nets or filamentous pseudopods for feeding) (Stechmann & Cavalier-Smith 2002, 2003; Cavalier-Smith 2004; Keeling 2004; Simpson & Roger 2004; Keeling *et al.* 2005).

Bikonts were defined as all eukaryotes ancestrally having two centrioles and cilia, with the anterior one being the younger and undergoing ciliary transformation to become the posterior cilium with a modified structure in its second cell cycle (Cavalier-Smith 2002). Unikonts were proposed to have had a last common ancestor with only one centriole and one cilium. It has long been known that many unikonts have two centrioles and some even two cilia but these were considered derived complications. When unikonts have two cilia, the anterior one never transforms into the posterior one. As many bikonts are secondarily uniciliate, the unikont/bikont distinction stresses fundamental differences in centriolar development and inferred ancestral state, not the number of centrioles or cilia per cell, which is evolutionarily more labile. Based on a rare gene fusion and other molecular cladistic characters, as well as basic differences in microtubular cytoskeleton and ciliary development (Cavalier-Smith 2002), the root of the eukaryote tree of life was proposed to lie between bikonts and unikonts (Stechmann & Cavalier-Smith 2002, 2003; Richards & Cavalier-Smith 2005). All recent multigene trees (e.g. Burki *et al.* 2007; Rodríguez-Ezpeleta *et al.* 2007) strongly support a bipartition of eukaryotes into unikonts and bikonts and are compatible with the root lying between

\* Author for correspondence (k.s.jakobsen@bio.uio.no).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2008.1358> or via <http://journals.royalsociety.org>.

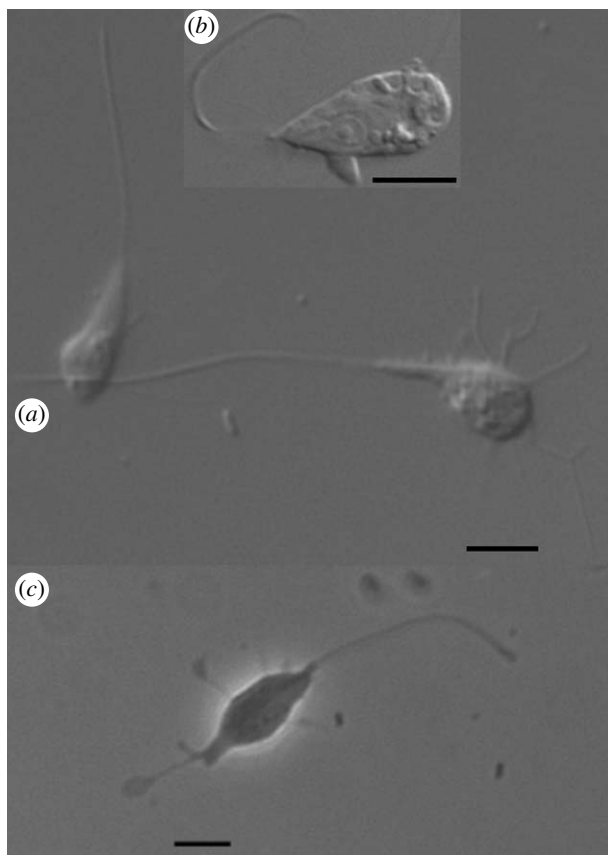


Figure 1. *In vivo* morphology of *B. anathema*. Light micrographs of unstained living *B. anathema* cells. (a) 400 $\times$  DIC image highlighting the numerous branching pseudopodia and widened cell sheath at the base of the single flagellum. (b) Inset 630 $\times$  DIC image showing the position of the nucleus containing a centrally located nucleolus. (c) 400 $\times$  phase-contrast image highlighting the flattened pseudopodial attachments to the substrate. Scale bars, 5  $\mu$ m.

them, though a recent paper on just a few genes raises a potential problem for the simplest interpretation of these data (Kim *et al.* 2006). To test it more thoroughly and better eliminate alternatives, additional putatively derived cladistic characters need to be identified (Rodríguez-Ezpeleta *et al.* 2007), and other little studied lineages must be included in multigene analyses.

We focus here on the phylogenomics of one such key lineage, the breviate amoeboflagellates (Cavalier-Smith *et al.* 2004)—a group that has defied placement in either unikonts or bikonts or any of the six eukaryotic supergroups, and whose correct placement is likely to illuminate the primary eukaryotic divergence.

*Breviata anathema* (previously misidentified as *Mastigamoeba invertens*) is a deeply branching anaerobic amoeboflagellate eukaryote, which has been notoriously difficult to place phylogenetically (Cavalier-Smith *et al.* 2004; Walker *et al.* 2006), and has some apparent morphological affinities with unikonts (i.e. its amoeboid cell body and single flagellum) and some with bikonts (two basal bodies); its filose pseudopodia (micrographs, figure 1) differ from those of either group. In single-gene phylogenetic analyses of the small subunit ribosomal RNA gene (18S) and the largest subunit of DNA-dependent RNA polymerase II (RPB1), the position of *Breviata* is very unstable; it variably associates with the excavates, apusomonads (themselves either excavates or still earlier

branching bikonts) and/or planomonads (formerly misidentified as *Ancyromonas*; see Cavalier-Smith *et al.* 2008) or with Amoebozoa, but no position is significantly supported (Bolivar *et al.* 2001; Cavalier-Smith *et al.* 2004; Walker *et al.* 2006). Amoebozoa, the group to which we now show *Breviata* belongs, is probably one of the earliest branches from the eukaryotic cenancestor and important for deducing its characteristics (Cavalier-Smith 2002; Richards & Cavalier-Smith 2005). Although the name Amoebozoa is old (Lühe 1913), it has only recently been recognized as a phylogenetically coherent group, with many unrelated amoebae now being excluded (Cavalier-Smith 1998; Cavalier-Smith & Chao 2003) and its classification revised (Cavalier-Smith *et al.* 2004; Nikolaev *et al.* 2006). Amoebozoa currently include classical naked and testate lobose amoebae, anaerobic Archamoebae (Entamoebae and pelobionts) and mycetozoon slime moulds (Cavalier-Smith *et al.* 2004), but exclude all amoeboid protozoa with true filopodia (ones that draw the cell forward by contraction), which instead belong to the bikont phylum Cercozoa that includes the chlorarachnean algae (Cavalier-Smith & Chao 2003). However, based on phylogenetic analyses and ultrastructural features, Cavalier-Smith *et al.* (2004) proposed a new class Breviatea including *Breviata* and two environmental sequences that clustered together with *Breviata* in 18S rRNA phylogenies, and postulated breviate as the out-group to all other Amoebozoa.

As multigene analyses usually generate more robust phylogenetic inferences than single genes (Bapteste *et al.* 2002; Burki *et al.* 2007), we constructed a cDNA library from *B. anathema* and sequenced approximately 4100 clones and reconstructed global eukaryote phylogeny using approximately 17 300 amino acid characters (figure 2). We also searched our database for mitochondria-related genes, as *Breviata* is also of special evolutionary interest as an anaerobic/microaerophilic organism with unusual hydrogenosome-like organelles, whose putative mitochondrial nature is controversial (Walker *et al.* 2006). As is well known, several eukaryote lineages within fungi, Amoebozoa (pelobionts, *Entamoeba*), ciliates, heterokonts (*Blastocystis*) and excavates (Heterolobosea, Preaxostyla, parabasalids, diplomonads and *Carpodimonas*) independently modified their mitochondria into anaerobic energy-generating organelles (hydrogenosomes) or the more degenerate mitosomes (Tielens *et al.* 2002; van der Giezen & Tovar 2005; Barbera *et al.* 2007). Since all groups other than breviate that putatively represented descendants of a pre-mitochondrial eukaryotic lineage have now been investigated and shown to contain mitochondrial-related remnants (i.e. organelles or genes) (Hampl *et al.* 2008), the only remaining known lineage that might be ancestrally amitochondriate is the breviate.

However, genes that trace their ancestry to the mitochondrion clearly demonstrate a mitochondrial history for *Breviata*.

## 2. MATERIAL AND METHODS

### (a) Library construction and EST sequencing

*B. anathema* (strain ATCC 50338) was cultured with one or two unidentified bacteria as food in tightly sealed 500 ml tissue culture flasks containing 75 ml ATCC 1773 medium at room temperature (approx. 21°C). Total RNA was isolated from cells harvested by centrifugation using Tri reagent

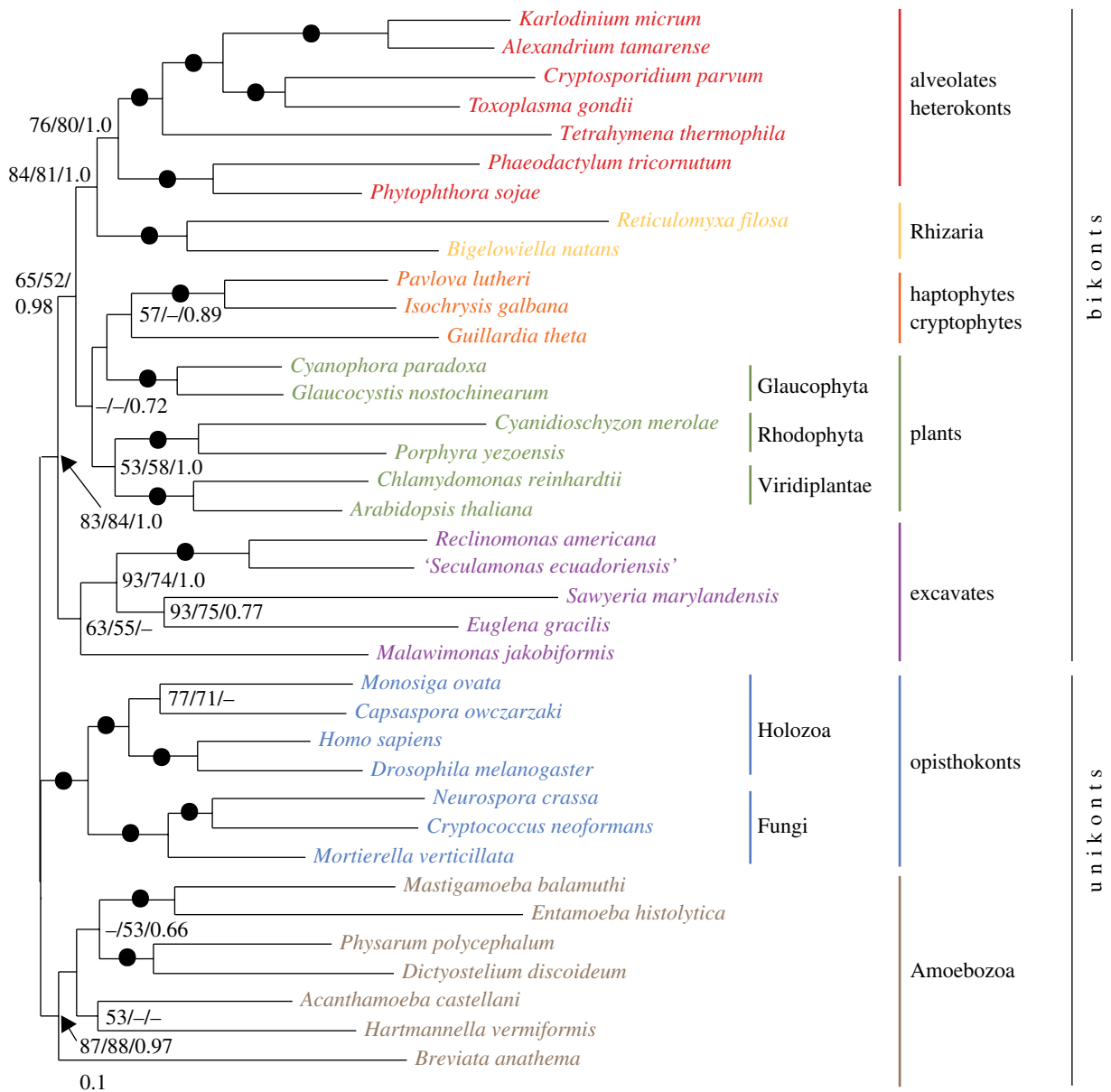


Figure 2. A global phylogeny of eukaryotes. Maximum-likelihood tree with bootstrap support values (BV) from an amino acid alignment of 78 concatenated genes (17 283 characters) inferred using RAXML and TREEFINDER (both giving identical topology; the RAXML tree is shown). Bayesian PP support values for bipartitions are also shown if more than 0.50. Filled circles denote support values of 100% BV and 1.0 PP, and dash (–) denotes support value below 50% BV or 0.50 PP. Nodes without denotation received less than 50% BV and less than 0.50 PP.

(Sigma-Aldrich, St Louis, MO, USA). A non-normalized, directional, ‘microquantity cDNA library’ was constructed in the plasmid vector pAGEN-1 by Agencourt Bioscience, Corp. (Beverly, MA, USA). Approximately, 4100 randomly picked clones were 5'-end sequenced; the EST sequences were subsequently quality checked and assembled to contigs using a Phred/Phrap pipeline at the freely available Bioportal service at the University of Oslo (<http://www.biportal.uio.no>).

#### (b) Multigene alignment construction

BLASTx analyses (<http://www.ncbi.nlm.nih.gov/BLAST>) of *Breviata* singletons and contigs were performed to identify gene similarities. *Breviata* sequences and significant hits ( $E\text{-value} > 1e^{-5}$ ) from a range of other publicly available sequences from different databases (TBestDB, <http://tbestdb.bcm.umontreal.ca/searches/login.php>; NCBIest and NCBIInr database) were added to the existing single-gene alignments (Rodríguez-Ezpeleta et al. 2005; Burki et al. 2007).

Ambiguously aligned characters were selected manually and excluded from the analyses. For each single-gene alignment, orthologous gene copies were identified by manual inspection of phylogenetic trees and bootstrap values (BV) inferred with PhyML (rtREV substitution model, 100 bootstrap replicates; Guindon & Gascuel 2003). Additionally, for taxa with two or more nearly identical sequences, the sequence displaying the shortest branch length on the tree was kept. The final multigene dataset contained 78 genes (17 280 amino acid characters) and 37 taxa. Taxa sampled were chosen to reflect the evolutionary range of eukaryotes, and the genes selected are based on the genes detected in the *Breviata* library. Details about taxon sampling and genes used in the analyses are given in table S1 in the electronic supplementary material.

Three fast-evolving excavates were excluded from the main analyses owing to their long branches (Simpson et al. 2006), known to cause long-branch attraction artefacts in phylogenetic trees (Philippe 2000), but were included in an

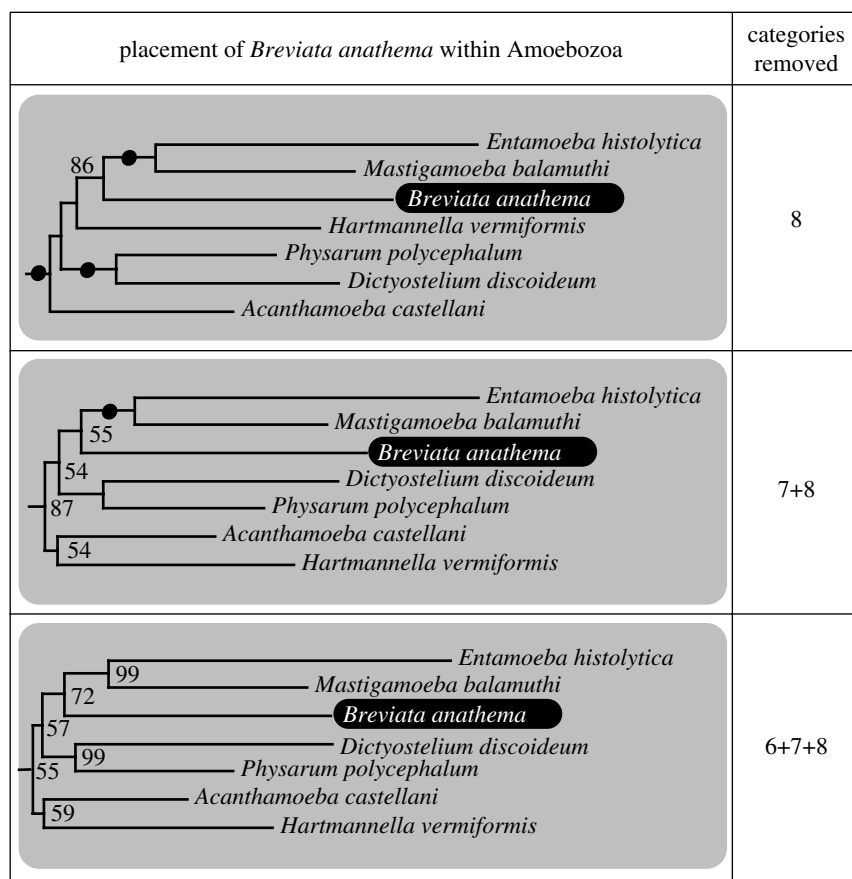


Figure 3. The placement of *Breviata* within Amoebozoa in three maximum-likelihood phylogenies with BV inferred with RAXML after removing categories of fast-evolving sites. Only the Amoebozoa branch is shown and global trees are shown in figure S2 in the electronic supplementary material. Categories 6, 7 and 8 refer to the sites removed; category 8 comprises the fastest evolving sites. Filled circles denote support values of 100% BV.

additional analysis shown in supplementary material (see figure S1 in the electronic supplementary material).

The impact of fast-evolving sites on the phylogeny was assessed by estimation with codonML in PAML (Yang 2007) under eight rate categories and subsequent site removal script applied to the alignment (S. Kumar, Å. Skjæveland, T. Ruden, A. Botnen & K. Shalchian-Tabrizi 2008, unpublished data). ML bootstrap consensus trees were inferred (as described below) from 100 pseudoreplicate datasets after the three fastest site-rate categories were removed (see figure S2 in the electronic supplementary material). Support for Amoebozoa and for the position of *Breviata* in optimal trees is shown in figure 3.

### (c) Phylogenetic analyses and approximately unbiased test

All phylogenetic analyses were performed on the Biportal at the University of Oslo (<http://www.biportal.uio.no>). Maximum-likelihood phylogeny of the concatenated data was inferred with RAXML MPI v. 2.2.3 (Stamatakis 2006) and TREEFINDER (Jobb et al. 2004). The rtREV+F evolutionary model was preferred by PROTTEST v. 1.3 under the Akaike information criterion with four GAMMA rate categories (Posada & Crandall 1998). Topological tree searches were performed with 100 randomly generated starting trees, while bootstrap analysis was performed on 100 pseudoreplicates and one random starting tree for each replicate, with the same evolutionary model as the initial search. In the RAXML, analyses trees were inferred under PROTMIX (Stamatakis 2006).

Bayesian inference used PHYLOBAYES v. 2.3 (Lartillot & Philippe 2004), with the CAT evolutionary model, a gamma-distributed across-site variation (four discrete rate categories) and random starting tree. Changes in log likelihood as a function of time were used to estimate whether the two parallel chains had reached a stationary state. This was then used to set the burn-in and compare the frequency of the bipartitions between several independent runs. The largest discrepancy (maxdiff) between the bipartitions was less than 0.1, and therefore we considered the Markov chain Monte Carlo chains to have converged. The tree and PP values presented in figure 2 are a consensus of the cold chains from the two independent runs.

The approximately unbiased (AU) tests were performed on the dataset that included all sites and on datasets with categories of fast-evolving sites removed (see table S2 in the electronic supplementary material). Site likelihoods were calculated in RAXML and the AU test performed with CONSEL (Shimodaira & Hasegawa 2001) using the rtREV evolutionary model, default scaling and replicate values.

## 3. RESULTS AND DISCUSSION

### (a) A global phylogeny including *B. anathema*

In our phylogeny, *Breviata* is convincingly placed with Amoebozoa (supported with 87/88% BV and 0.97 PP value; figure 2) by both maximum-likelihood (inferred with RAXML and TREEFINDER, respectively) and Bayesian methods. Removing the fastest evolving sites of the alignment did not influence this placement (figure 3; see

figure S2 in the electronic supplementary material). Removing the fastest evolving sites increased the bootstrap support to 100 per cent BV for *Breviata* grouping with Amoebozoa (figure 3a). Sequential removal of additional fast-site categories decreased the support for most supergroups, including Amoebozoa, but the relationship of *Breviata* + Amoebozoa was always recovered. In all trees with fastest evolving sites removed, the clear-cut separation into unikonts and bikonts (with *Breviata* among the unikonts) was even more strongly supported than that shown in figure 2 (88, 97, 95% BV; see figure S2 in the electronic supplementary material). An additional phylogeny including three additional fast-evolving excavate taxa (*Giardia intestinalis*, *Trichomonas vaginalis* and *Trimastix pyriformis*; see figure S1 in the electronic supplementary material) also supported the placement of *Breviata* with Amoebozoa, but somewhat less strongly. Hence, this relationship is robust and not sensitive to the removal of fast-evolving sites or to taxon sampling. The alternative placement of *Breviata* within bikonts suggested by many single-gene trees (Cavalier-Smith *et al.* 2004; Shalchian-Tabrizi *et al.* 2006; Walker *et al.* 2006) is not seen in any inferred multigene trees, and this topology was rejected by the AU tests of the reduced datasets from which the fastest evolving sites were successively removed (AU test; see table S2 in the electronic supplementary material).

Although grouping of *Breviata* with Amoebozoa is strong, bootstrap support for placing *Breviata* as a sister to—rather than among—the other amoebozoan taxa is weak. Accordingly, the AU tests did not reject the possibility that *Breviata* may branch among other Amoebozoa as sister to the other anaerobic amoebae (Archamoebae: *Entamoeba* and *Mastigamoeba*; see table S2 in the electronic supplementary material) and this sister relationship is supported in two of the trees inferred after removing fast-evolving sites (figure 3a,c). However, it is more likely that *Breviata* is sister to the other Amoebozoa, owing to its lack of four sequence signatures in the 18S rRNA gene that other Amoebozoa all share; single nucleotide substitutions at positions 385, 777 and 1010 and a 1–2 nucleotide insertion in the loop between positions 1060 and 1064 (Fahrni *et al.* 2003). If *Breviatea* were sisters to Archamoebae, all four signatures must have reverted to the ancestral state found in all out-groups to Amoebozoa (Fahrni *et al.* 2003), which is unlikely as most other Amoebozoa have all four of these signatures, and all have at least two (Fahrni *et al.* 2003).

Overall, our inferred phylogeny (figure 2) is congruent with other recent global eukaryotic phylogenies (Burki *et al.* 2007; Rodríguez-Ezpeleta *et al.* 2007). Several lineages are strongly supported by maximum-likelihood BV and PP values, including Holozoa (animals + choanoflagellates), fungi, opisthokonts, Rhodophyta, Glaucophyta, Viridiplantae, Haptophyta, Alveolata, Rhizaria and Heterokonta. Our tree is congruent with several higher order relationships with BV values above 80 per cent: Amoebozoa, including *Breviata* (87/88% BV, 0.99 PP) and a grouping of alveolates, heterokonts (stramenopiles) and Rhizaria—the putative SAR assembly, noted previously in several recent phylogenies (84/81% BV; 0.97 PP) (Burki *et al.* 2007; Hackett *et al.* 2007; Rodríguez-Ezpeleta *et al.* 2007). The putative basal bifurcation between unikonts and bikonts is supported by 83/84 per cent BV (1.00 PP). Excavates, excluding Preaxostyla (*Trimastix* +

oxymonads), Eopharyngia (diplomonads and retortamonads) and parabasalids, are monophyletic but with weak bootstrap support (63/55% BV). This clade is not recovered in the Bayesian phylogeny. Plantae are paraphyletic here owing to the inclusion of haptophytes and cryptomonads.

A minority of 18S rRNA analyses have suggested a specific affiliation of *Breviata* to apusomonads (Walker *et al.* 2006), but too few protein-coding genes are available from apusomonads for us to test this hypothesis directly. Likewise, the phylogenetic position of apusomonads is controversial, with ultrastructural and gene fusion evidence suggesting a bikont affinity (Karpov & Zhukov 1986; Stechmann & Cavalier-Smith 2002) while two- to six-gene phylogenies place *Apusomonas proboscidea* as sister to opisthokonts (Kim *et al.* 2006). However, when  $\alpha$ -tubulin was excluded from the multigene analyses of Kim *et al.* (2006), the placement of *A. proboscidea* as sister to Amoebozoa could not be rejected (Kim *et al.* 2006). Thus, there is no evidence suggesting that *Breviata* is misplaced in our tree.

#### (b) *Breviate amoebae are unusual amoebozoans*

In all our multigene trees, *Breviata* is placed with Amoebozoa with high support. The precise placement within the group, however, is not consistent in the trees inferred, as some of them support a sister relationship between Archamoebae and *Breviata*, while others indicate that *Breviata* is sister to the remaining Amoebozoa (figures 2 and 3). Notably, the absence of the Amoebozoa-specific substitutions in the 18S sequence indicates that the latter hypothesis, consistent with the hypothesis proposed by Cavalier-Smith *et al.* (2004), is more likely. Walker *et al.* (2006) reasonably argued that because *Breviata* is not closely similar in morphology to any of the other classes of ciliated Amoebozoa it does not belong in any of them (Walker *et al.* 2006). However, their conclusion that it is therefore not an amoebozoan did not take into account the possibility of a common ancestry plus later substantial morphological divergence from the other classes, which now appears to be the case. Indeed, amoebozoan morphological diversity has been expanded by careful observations that reveal a unique gait in *Breviata* locomotion. These amoebae travel by ‘walking’ with thin but robust leg-like pseudopodia that emanate from the anterior of the cell body, and adhere to the substratum, while the cell body proceeds forward just as a package travelling on a roller conveyor or ‘tractor on treads’ (figure 1). The filose ‘legs’ often remain as trailing filaments before they retract into the cell body. This character distinguishes *Breviata* from other organisms, as no other eukaryote has even vaguely similar motor movements.

Prior to the addition of *Breviatea*, Amoebozoa comprised two well-defined subphyla: the often ciliated Conosa (Mycetozoa, Archamoebae) characterized by a conical microtubular skeleton diverging from the centriole or centrosome, and the purely amoeboid Lobosa that lack cilia, centrioles and cytoplasmic microtubules (Cavalier-Smith 1998). Our demonstration that *Breviata* is an amoebozoan significantly increases the cellular diversity of the phylum owing to its unusual pseudopodial morphology, mode of locomotion and rather complex cytoskeleton. In marked contrast to the also anaerobic

Archamoebae, *Breviata* has two centrioles and a substantially more asymmetric microtubular cytoskeleton. These differences, plus the presence of Golgi stacks in *Breviata*, but not Archamoebae, justify their being in separate classes (Cavalier-Smith *et al.* 2004; as do the four contrasting rRNA signatures mentioned above), but (contrary to Walker *et al.* 2006) are not enough to merit separate phyla. Thus, there are now three broadly different cytoskeletal patterns in Amoebozoa.

### (c) *Implications for ultrastructural evolution in early eukaryotes*

The ancestral cellular structure for Amoebozoa was argued to be a uniciliate, unicentriolar amoeba with a radially symmetric pericentriolar microtubular cone (Cavalier-Smith *et al.* 2004). However, as the uniciliate *Breviata* possesses two centrioles, one of which serves as the basal body of the cilium resulting in an asymmetric cytoskeleton (Walker *et al.* 2006), this interpretation needs some re-evaluation. As there are also other amoebozoan lineages with two basal bodies, such as myxogastriids and a few protostelids, the two basal bodies in *Breviata* do not contradict the inference that *Breviata* is an amoebozoan, but merely suggest that it is not an Archamoeba (Cavalier-Smith *et al.* 2004). If *Breviata* were sister to Archamoebae, as some trees excluding faster evolving sites suggest but which the rRNA signatures render unlikely, one could argue more strongly that its having a second barren centriole is a derived state. However, our more inclusive trees and 18S rRNA signatures in combination indicate that *Breviata* is probably sister to all previously accepted Amoebozoa. This makes it harder to infer the ancestral state of Amoebozoa, in which there are now two groups with two centrioles/basal bodies (*Breviata*, myxogastriids), three with one centriole per kinetid (*Multicilia*, *Phalansterium*, Archamoebae) and one with a mixture (protostelids). Thus, a double centriolar ancestral state for Amoebozoa is almost as parsimonious as the single centriolar scenario (Cavalier-Smith *et al.* 2004), especially as deeply branching opisthokonts (chytids and choanoflagellates), the sister group to Amoebozoa, have two centrioles. With respect to the cytoskeleton, the marked asymmetry found in *B. anathema* contrasts with the hypothesized symmetrical ancestral state of Amoebozoa (Cavalier-Smith 2002). This asymmetry could be secondarily derived in *B. anathema* and does not imply an affinity to the asymmetric bikonts since the detailed arrangement of their ciliary roots differ substantially. Thus, the inclusion of *Breviata* within Amoebozoa as its most divergent group has important implications for the ultrastructural evolution and likely ancestral state of the cytoskeleton and centrioles in Amoebozoa and eukaryotes generally. Our findings make it important to study both the cytoskeleton and the pattern of ciliary and centriolar development more thoroughly in *B. anathema* and test their generality among different breviate. As contrasting modes of ciliary development were a key aspect of the original recognition of the primary dichotomy between bikont and unikont eukaryotes (Cavalier-Smith 2002), such studies are of key significance for clarifying the basic organization of the earliest eukaryote cells. Unfortunately, ciliary development is unstudied for *Breviata* and for apusomonads, whose putative inclusion within unikonts (Kim *et al.* 2006), is unexpected, given their biciliate (not

necessarily bikont; for the distinction see Cavalier-Smith 2002) nature and the structure of their ciliary roots (Molina & Nerad 1991), and needs further confirmation by multigene analyses.

### (d) *The mitochondria-like organelle in B. anathema was probably derived independently from the other anaerobic lineages*

In our *Breviata* cDNA library, we identified key mitochondria-derived nuclear-encoded genes often seen in amitochondrial taxa that trace their ancestry back to an  $\alpha$ -proteobacterial ancestor (here shown by *cpn60* (see figure S3 in the electronic supplementary material) and *tim17* (data not shown)). This clearly rejects the possibility that *Breviata* is a pre-mitochondrial eukaryote, and suggests that the dense organelles bounded by two membranes seen proximal to the nucleus in *Breviata* are mitochondria-related organelles. Further investigations of mitochondrial function in *Breviata*, including a search for hydrogenase and biochemical studies, are now needed. If *Breviata* is sister to other amoebozoa, the anaerobic adaption of the mitochondria in *Breviata* occurred independently of other known cases. However, our multigene trees and AU tests do not exclude the possibility that Archamoebae and *Breviata* form a single secondarily anaerobic amoebozoan clade.

All extant eukaryotes examined in detail, even anaerobic 'amitochondriate' eukaryotes, have nuclear genes whose phylogenetic history is best explained by entry into the eukaryote lineage with the mitochondrion endosymbiont. It is thus unlikely that the anaerobic nature of *Breviata* represents the ancestral state of Amoebozoa, even though our data suggest that *Breviata* may be the deepest diverging amoebozoan lineage. The ancestral amoebozoan must have been at least facultatively aerobic, though it could have been a facultative aerobe/anaerobe, as many have postulated for the ancestral eukaryote (Cavalier-Smith 2006). Possibly aerobic members of the *Breviata* clade will be discovered.

### (e) *Phylogenomics of unassigned species resolves key questions in eukaryote evolution*

The challenging task of resolving eukaryotic global phylogeny has progressed through phylogenomic analysis of major lineages (e.g. Nikolaev *et al.* 2004; Rodríguez-Ezpeleta *et al.* 2005; Burki & Pawlowski 2006; Burki *et al.* 2007; Patron *et al.* 2007; Rodríguez-Ezpeleta *et al.* 2007). Here, we demonstrated that investigating single, deeply diverging nomadic species is also crucial for improving our understanding of early evolutionary history of major lineages of eukaryotes. Placing the previously unaffiliated breviate, with their unique cytoskeletal pattern, in a clade with other Amoebozoa illuminates the evolutionary diversity of Amoebozoa and raises new questions concerning the nature of ancestral amoebozoan and of the unikont–bikont bifurcation suspected to reside at the base of the eukaryote tree.

We thank Cédric Berney for helpful comments on the manuscript, Dag Klaveness for fruitful discussions and Surendra Kumar for the site removal script. The Norwegian Research Council has granted scholarships to K.S.-T., M.A.M., R.J.S.O. and research project to K.S.J. T.C.-S. thanks NERC and the Canadian Institute for Advanced Research Evolutionary Biology Program for fellowship support and NERC for research grants. The pan-Canadian

collaboration Protist EST Program (PEP: <http://megasun.bch.umontreal.ca/pepdb/pep.html>) generated sequence data for some of the species included in the phylogenetic analyses.

## REFERENCES

- Baldauf, S. L. 2003 The deep roots of eukaryotes. *Science* **300**, 1703–1706. (doi:10.1126/science.1085544)
- Bapteste, E. et al. 2002 The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba*, and *Mastigamoeba*. *Proc. Natl Acad. Sci. USA* **99**, 1414–1419. (doi:10.1073/pnas.032662799)
- Barbera, M. J., Ruiz-Trillo, I., Leigh, J., Hug, L. A. & Roger, A. J. 2007 The diversity of mitochondrion-related organelles amongst eukaryotic microbes. In *Origin of mitochondria and hydrogenosomes* (eds W. Martin & M. Müller), pp. 239–268. Heidelberg, Germany: Springer.
- Bolivar, I., Fahrni, J. F., Smirnov, A. & Pawlowski, J. 2001 SSU rRNA-based phylogenetic position of the genera *Amoeba* and *Chaos* (Lobosea, Gymnamoebia): the origin of gymnamoebae revisited. *Mol. Biol. Evol.* **18**, 2306–2314.
- Burki, F. & Pawlowski, J. 2006 Monophyly of Rhizaria and multigene phylogeny of unicellular bikonts. *Mol. Biol. Evol.* **23**, 1922–1930. (doi:10.1093/molbev/msl055)
- Burki, F., Shalchian-Tabrizi, K., Minge, M. A., Skjæveland, Å., Nikolaev, S. I., Jakobsen, K. S. & Pawlowski, J. 2007 Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE* **2**, e790. (doi:10.1371/journal.pone.0000790)
- Cavalier-Smith, T. 1998 A revised six-kingdom system of life. *Biol. Rev.* **73**, 203–266. (doi:10.1017/S0006323198005167)
- Cavalier-Smith, T. 2002 The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int. J. Syst. Evol. Microbiol.* **52**, 297–354.
- Cavalier-Smith, T. 2004 Only six kingdoms of life. *Proc. R. Soc. B* **271**, 1251–1262. (doi:10.1098/rspb.2004.2705)
- Cavalier-Smith, T. 2006 Origin of mitochondria by enslavement of a photosynthetic purple bacterium. *Proc. R. Soc. B* **273**, 1943–1952. (doi:10.1098/rspb.2006.3531)
- Cavalier-Smith, T. & Chao, E. E. Y. 2003 Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist* **154**, 341–358. (doi:10.1078/143446103322454112)
- Cavalier-Smith, T., Chao, E. E. Y. & Oates, B. 2004 Molecular phylogeny of Amoebozoa and the evolutionary significance of the unikont *Phalansterium*. *Eur. J. Protistol.* **40**, 21–48. (doi:10.1016/j.ejop.2003.10.001)
- Cavalier-Smith, T., Chao, E. E., Stechmann, A., Oates, B. & Nikolaev, S. 2008 Planomonadida ord. nov. (Apusozoa): ultrastructural affinity with *Micronuclearia podoventralis* and deep divergences within *Planomonas* gen. nov. *Protist* **159**, 535–562.
- Corliss, J. O. 1984 The kingdom Protista and its 45 phyla. *BioSystems* **17**, 87–126. (doi:10.1016/0303-2647(84)90003-0)
- Fahrni, J., Bolivar, I., Berney, C., Nassonova, E., Smirnov, A. & Pawlowski, J. 2003 Phylogeny of lobose amoeba based on actin and small-subunit ribosomal RNA genes. *Mol. Biol. Evol.* **20**, 1881–1886. (doi:10.1093/molbev/msg201)
- Guindon, S. & Gascuel, O. 2003 A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* **52**, 696–704. (doi:10.1080/10635150390235520)
- Hackett, J. D., Yoon, H. S., Li, S., Reyes-Prieto, A., Rummele, S. E. & Bhattacharya, D. 2007 Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of Rhizaria with chromalveolates. *Mol. Biol. Evol.* **24**, 1702–1713. (doi:10.1093/molbev/msm089)
- Hampl, V., Silberman, J. D., Stechmann, A., Diaz-Trivino, S., Johnson, P. J. & Roger, A. J. 2008 Genetic evidence for a mitochondriate ancestry in the ‘amitochondriate’ flagellate *Trimastix pyriformis*. *PLoS ONE* **3**, e1383. (doi:10.1371/journal.pone.0001383)
- Jobb, G., von Haeseler, A. & Strimmer, K. 2004 TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4**. (doi:10.1186/1471-2148-4-18)
- Karpov, S. A. & Zhukov, B. F. 1986 Ultrastructure and taxonomic position of *Apusomonas proboscidea* Alexeieff. *Arch. Protistenkd.* **131**, 13–26.
- Keeling, P. 2004 A brief history of plastids and their hosts. *Protist* **155**, 3–7. (doi:10.1078/1434461000156)
- Keeling, P. J., Burger, G., Durnford, D. G., Lang, B. F., Lee, R. W., Pearlman, R. E., Roger, A. J. & Gray, M. W. 2005 The tree of eukaryotes. *Trends. Ecol. Evol.* **20**, 670–676. (doi:10.1016/j.tree.2005.09.005)
- Kim, E., Simpson, A. G. B. & Graham, L. E. 2006 Evolutionary relationships of apusomonads inferred from taxon-rich analyses of 6 nuclear encoded genes. *Mol. Biol. Evol.* **23**, 2455–2466. (doi:10.1093/molbev/msl120)
- Lartillot, N. & Philippe, H. 2004 A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109. (doi:10.1093/molbev/msh112)
- Molina, F. I. & Nerad, T. A. 1991 Ultrastructure of *Amastigomonas bermudensis* ATCC-50234 sp. nov.—a new heterotrophic marine flagellate. *Eur. J. Protistol.* **27**, 386–396.
- Nikolaev, S. I., Berney, C., Fahrni, J. F., Bolivar, I., Polet, S., Mylnikov, A. P., Aleshin, V. V., Petrov, N. B. & Pawlowski, J. 2004 The twilight of Heliozoa and rise of Rhizaria, an emerging supergroup of amoeboid eukaryotes. *Proc. Natl Acad. Sci. USA* **101**, 8066–8071. (doi:10.1073/pnas.0308602101)
- Nikolaev, S. I., Berney, C., Petrov, N. B., Mylnikov, A. P., Fahrni, J. F. & Pawlowski, J. 2006 Phylogenetic position of *Multicilia marina* and the evolution of Amoebozoa. *Int. J. Syst. Evol. Microbiol.* **56**, 1449–1458. (doi:10.1099/ijs.0.63763-0)
- Patron, N. J., Inagaki, Y. & Keeling, P. J. 2007 Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr. Biol.* **17**, 887–891. (doi:10.1016/j.cub.2007.03.069)
- Philippe, H. 2000 Opinion: long branch attraction and protist phylogeny. *Protist* **151**, 307–316. (doi:10.1078/S1434-4610(04)70029-2)
- Posada, D. & Crandall, K. A. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. (doi:10.1093/bioinformatics/14.9.817)
- Richards, T. A. & Cavalier-Smith, T. 2005 Myosin domain evolution and the primary divergence of eukaryotes. *Nature* **436**, 1113–1118. (doi:10.1038/nature03949)
- Rodríguez-Ezpeleta, N. et al. 2005 Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr. Biol.* **15**, 1325–1330. (doi:10.1016/j.cub.2005.06.040)
- Rodríguez-Ezpeleta, N., Brinkmann, H., Burger, G., Roger, A. J., Gray, M. W., Philippe, H. & Lang, B. F. 2007 Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr. Biol.* **17**, 1420–1425. (doi:10.1016/j.cub.2007.07.036)
- Shalchian-Tabrizi, K. et al. 2006 Telonemia, a new protist phylum with affinity to chromist lineages. *Proc. R. Soc. B* **273**, 1833–1842. (doi:10.1098/rspb.2006.3515)
- Shimodaira, H. & Hasegawa, M. 2001 CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **16**, 296–297.
- Simpson, A. G. B., Inagaki, Y. & Roger, A. J. 2006 Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of ‘primitive’ eukaryotes. *Mol. Biol. Evol.* **23**, 615–625. (doi:10.1093/molbev/msj068)

- Simpson, A. G. B. & Roger, A. J. 2004 The real 'kingdoms' of eukaryotes. *Curr. Biol.* **14**, R693–R696. (doi:10.1016/j.cub.2004.08.038)
- Stamatakis, A. 2006 RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)
- Stechmann, A. & Cavalier-Smith, T. 2002 Rooting the eukaryote tree by using a derived gene fusion. *Science* **297**, 89–91. (doi:10.1126/science.1071196)
- Stechmann, A. & Cavalier-Smith, T. 2003 The root of the eukaryote tree pinpointed. *Curr. Biol.* **13**, R665–R666. (doi:10.1016/S0960-9822(03)00602-X)
- Tielens, A. G. M., Rotte, C., van Hellemond, J. J. & Martin, W. 2002 Mitochondria as we don't know them. *TiBS* **27**, 564–572. (doi:10.1016/S0968-0004(02)02193-X)
- van der Giezen, M. & Tovar, J. 2005 Degenerate mitochondria. *EMBO Rep.* **6**, 525–530. (doi:10.1038/sj.embor.7400440)
- Walker, G., Dacks, J. B. & Embley, T. M. 2006 Ultrastructural description of *Breviata anathema*, n. gen., n. sp., the organism previously studied as '*Mastigamoeba invertens*'. *J. Eukaryot. Microbiol.* **53**, 65–78. (doi:10.1111/j.1550-7408.2005.00087.x)
- Yang, Z. 2007 PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591. (doi:10.1093/molbev/msm088)